AF JFW
2655

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

First Named Inventor	: En-Dong Xun	Appeal No. ---
Appln. No.	: 09/757,836	
Filed	: January 10, 2001	Group Art Unit: 2655
For	: METHOD AND APPARATUS FOR PERFORMING MACHINE TRANSLATION USING A UNIFIED LANGUAGE MODEL AND TRANSLATION	Examiner: Michael A. Lewis
Docket No.:	M61.12-0342	

TRANSMITTAL OF APPEAL BRIEF
(PATENT APPLICATION - 37 C.F.R. §41.37)

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

I HEREBY CERTIFY THAT THIS PAPER IS BEING SENT BY U.S. MAIL, FIRST CLASS, TO THE COMMISSIONER FOR PATENTS, P.O. BOX 1450, ALEXANDRIA, VA 22313-1450, THIS
31st DAY OF August, 2005
PATENT ATTORNEY

Sir:

Transmitted herewith is the Appeal Brief ~~in~~ this application with respect to the Notice of Appeal filed on July 28, 2005.

FEE FOR FILING APPEAL BRIEF

Pursuant to 37 C.F.R. §41.20(b)(2) the fee for filing the Appeal Brief is $500.00.

The Director is authorized to charge any additional fees associated with this paper or credit any overpayment to Deposit Account No. 23-1123. A duplicate copy of this communication is enclosed.

Respectfully submitted,

WESTMAN, CHAMPLIN & KELLY, P.A.

The PTO ~~did not~~ receive the following listed item(s) as is burned

By: _____
Joseph R. Kelly, Reg. No. 34,847
Suite 1400 - International Centre
900 Second Avenue South
Minneapolis, Minnesota 55402-3319
Phone: (612) 334-3222 Fax: (612) 334-3312

JRK:slg

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

First Named Inventor	: En-Dong Xun	Appeal No. ---
Appln. No.	: 09/757,836	
Filed	: January 10, 2001	Group Art Unit: 2655
For	: METHOD AND APPARATUS FOR PERFORMING MACHINE TRANSLATION USING A UNIFIED LANGUAGE MODEL AND TRANSLATION	Examiner: Michael A. Lewis
Docket No.:	M61.12-0342	

BRIEF FOR APPELLANT

Mail Stop Appeal Brief-Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

I HEREBY CERTIFY THAT THIS PAPER IS BEING SENT BY U.S. MAIL, FIRST CLASS, TO THE COMMISSIONER FOR PATENTS, P.O. BOX 1450, ALEXANDRIA, VA 22313-1450, THIS
31st DAY OF August, 2005
Joseph R. Keating
PATENT ATTORNEY

Sir:

This is an appeal from an Office Action dated June 22, 2005, in which claims 1-16 were finally rejected. The Appellants respectfully submit that claims 1-16 are allowable, and request that the Board reverse the rejection of claims 1-16 and find that claims 1-16 are in condition for allowance.

09/06/2005 EFLORES 00000057 09757836
01 FC:1402 500.00 OP

Contents

REAL PARTY IN INTEREST

Microsoft Corporation, a corporation organized under the laws of the state of Washington, and having offices at One Microsoft Way, Redmond, Washington 98052, has acquired the entire right, title and interest in and to the invention, the application, and any and all patents to be obtained therefor, as set forth in the Assignment filed with the patent application and recorded on Reel 011749, frame 0340.

RELATED APPEALS AND INTERFERENCES

There are no known related appeals or interferences which will directly affect or be directly affected by or have a bearing on the Board's decision in this appeal.

STATUS OF THE CLAIMS

Claims 1-16 were originally presented. No claims have been added or canceled. Pending and rejected claims 1-16 are the subject of the present appeal.

STATUS OF AMENDMENTS

Appellants first amended the claims on October 4, 2004. Appellants filed an amendment on March 17, 2005, placing the claims in the form in which they were examined for the final Office Action dated June 22, 2005. No further amendment has been made.

SUMMARY OF CLAIMED SUBJECT MATTER

1. Introduction

The present invention relates to machine translation of languages, and specifically to the phrase translation of languages using a unified language model and translation model.

2. Brief Background

Machine translation involves a computer receiving input text either in written form, or in the form of speech, or in another suitable machine readable form. The computer uses a translation model in order to translate the words in the input text from a first language (in which they are input) to a second

language. Such prior systems have typically either used rules-based translators in which a set of rules is applied to the input text in order to arrive at a translation from the first language to the second language, or statistical translators in which statistical models are used to translate the words in the input from the first language to the second language.

3. The Present Invention

The present invention overcomes problems associated with prior approaches by using a model that calculates a transition probability based on a combination of things. First, a plurality of different possible linguistic patterns are identified based on the input text. A probability is calculated for each of the possible patterns, and a language model score is calculated which is indicative of the probability of a possible translation of the input text given the pattern. Similarly, a translation model score is calculated indicative of the probability of generating the input phrase given a possible translation of the input phrase in the linguistic pattern under analysis. A translation probability is calculated for the given linguistic pattern based on a combination of the language model probability for the pattern and the translation model probability for the pattern.

As a more concrete example, reference is now made to Figures 4A and 4B in the current specification. FIG. 4A shows a tree that represents an English language phrase (represented by the letter "E"). The nodes D and E on the tree in FIG. 4A are non-terminal nodes, while the nodes A, B and C represent terminal, or leaf nodes. The leaf nodes represent individual words in English phrase "E". It can be seen from FIG. 4A that the English phrase "E" is composed of a non-terminal phrase D and the English word C. The phrase D is composed of two English words A and B. See page 13, lines 18. et. seq.

FIG. 4B illustrates the variety of different linguistic

patterns that can be used in translating the phrase E. Those patterns are identified by numerals 300, 302, 304, 306, 308 and 310. Linguistic pattern 300 illustrates that the translation of phrase E can be formed by translation of the phrase D followed by the translation of the word C. Linguistic pattern 302 indicates that the translation of phrase E can be composed of a translation of the word C followed by a translation of the phrase D. Of course, since phrase D is actually made up of two words (words A and B) translation of phrase D can also be performed by translating the word A and following it with the translation of the word B or vise versa. This is indicated by patterns 304 and 306 in FIG. 4B. Patterns 308 and 310 show the same type of linguistic patterns, except that the expanded translation of phrase D follows translation of the word C.

Therefore, the bilingual pattern data store 210 (shown in FIG. 2) illustratively includes a plurality of English phrases (such as phrase E) followed by a corresponding plurality of linguistic patterns in the second language (such as the linguistic patterns set out in FIG. 4B) which correspond to, and are possible linguistic translation patterns of, the English phrase E. See page 13, lines 18-page 14, line 23.

Equation 5 on page 18 in the specification illustrates one embodiment in which the present invention can be used to generate a score corresponding to each of the possible linguistic patterns that, themselves, correspond to a possible translation. For this example, assume that the input text is in English and the output text is in Chinese. Therefore, in order to calculate a score corresponding to a possible translation, one of the linguistic patterns associated with that translation is selected and the probability of that pattern is calculated. This is referred to as the pattern probability, and it is combined with a language model score for the possible translation in the Chinese language, and a translation model score which is indicative of a

probability of an English language word, given the Chinese translation.

Because the pattern probability is factored into the score for a given possible translation, both contextual and structural information from the input text are retained in determining the score. This is highly advantageous over prior systems which were either limited by a rules-based translator, or which retained essentially no structural information in a statistical translator.

GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

1. Whether the specification fails to comply with the written description requirement under 35 U.S.C.§112, first paragraph.

2. Whether claims 1-16 are unpatentable under 35 U.S.C.§102(b) as being anticipated by Berger et al. US Patent No. 5,510,981.

The Appellants respectfully submit that claims 1-16 have been fully described in the specification in such a way as to reasonably convey to one skilled in the art that the inventors, at the time the application was filed, had possession of the claimed invention. The Appellants also respectfully submit that claims 1-16 are not anticipated by Berger et al., US Patent No. 5,510,981.

The Appellants thus respectfully request that the Board reverse the Examiner as to both grounds of rejection.

GROUPING OF CLAIMS

The claims do not stand or fall together, but are grouped as follows, and each group is believed to be independently patentable:

I. Claim 1;

II. Claims 6, 7, 12, 14 and 15;

III. Claim 2, 3, 9, 10 and 16;

IV. Claim 4, 11 and 13;

V. Claims 5 and 8;

ARGUMENT

1. The specification meets the written description requirement set out in 35 U.S.C. §112, first paragraph.

The Examiner rejected claims 1-16 under 35 U.S.C.§112, first paragraph, as failing to comply with the written description requirement. Specifically, the Examiner took issue with the language added in the amendment filed March 17, 2005, and in particular the language "wherein each of the plurality of possible linguistic patterns represents a grouping of components relative to the phrase[.]" The Examiner asserted that this was not described in the specification. Appellants respectfully request reversal of this rejection.

35 U.S.C. §112, first paragraph states, in pertinent part, as follows:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains…to make and use the same…

The present specification clearly meets this requirement with respect to the language added by amendment of March 17, 2005. Again, that language specifically states that "each of the plurality of possible linguistic patterns represents a grouping of components relative to the phrase [.]"

> The specification states that FIG. 4A shows:
> a tree for an English phrase (represented by "E"). The nodes D and E on the tree in FIG. 4A are non-terminal nodes, while the nodes A, B and C represent terminal, or leaf nodes, and thus, represent the individual words in phrase E. It can be seen from FIG. 4A that phrase E is composed

of a non-terminal phrase D and the English word C. The phrase D is composed of two English words A and B.

FIG. 4B illustrates the wide variety of linguistic patterns that can be used in translating the phrase E. Those phrases are identified by numerals 300, 302, 304, 306, 308 and 310. Linguistic pattern 300 illustrates that the translation of phrase E can be formed by translating the phrase D followed by a translation of the word C. Linguistic pattern 302 indicates that the translation of phrase E can be composed of a translation of the word C followed by a translation of the phrase D. Of course, since phrase D is actually made up of two words (A and B) translation of phrase D can also be performed by translating the word A and following it with the translation of the word B, or vice versa. This is indicated by patterns 304 and 306. Patterns 308 and 310 show the same type of linguistic patterns, except that the expanded translation of the phrase D follows translation of the word C.

Therefore, bi-lingual pattern data store 210 illustratively includes a plurality of English phrases (such as phrase E) followed by a corresponding plurality of linguistic patterns in the second language (such as the linguistic patterns set out in FIG. 4B) which correspond to, and are possible linguistic translation patterns of, the English phrase E. Page 13, line 18 to page 14, line 23 (emphasis added).

From the cited text, it is clear that the components of phrase E are phrase D and words A, B and C. It is also clear that non-terminal node D (which is a phrase) represents a grouping of leaf nodes (which are words) A and B. See specifically page 13, lines 26 and 27 which state "The phrase D is composed of the two English words A and B."

It is also clear that the linguistic patterns set out in FIG. 4B show translation of different groupings of the components of English phrase "E". For instance, pattern 300 "illustrates that the translation of phrase E can be formed by translating the phrase D followed by a translation of the word C." See page 14, lines 3-5. Similarly, pattern 304 represents translation of a different grouping of the component of phrase "E". "Since phrase D is actually made up of two words (A and B) translation of phrase D can also be performed by translating the word A and following it with the translation of the word B, or vice versa. This is indicated by patterns 304 and 306." See page 14, lines 8-12. Thus, instead of translating the phrase D as a whole, it can be translated by translating its individual words.

Clearly, then, each linguistic pattern represents "a grouping of components relative to the phrase." Applicant thus submits that the language submitted in the amendment filed March 17, 2005 is fully supported by both the specification and the drawings. Appellants respectfully request that the Board reverse the rejection under 35 U.S.C. §112, first paragraph.

2. Claim 1 is allowable over Berger et al.

Independent claim 1 is directed to a computer implemented method that specifically includes "identifying a plurality of possible linguistic patterns in the second language associated with the phrase in the first language, wherein each of the plurality of possible linguistic patterns represents a grouping of components relative to the phrase; and for each

pattern, calculating a translation probability for the pattern based on a combination of a language model probability for the pattern and a translation model probability for the pattern." This is simply neither taught nor suggested by Berger et al.

As described in section 1 above, the linguistic patterns represent a grouping of components relative to a phrase that is being translated. Berger et al. describes a translation environment in which two target hypotheses are generated wherein each target hypothesis is "a series of words." A match score is generated for each target hypothesis that includes a language model score and a word match score. Of course, as is well known in the art, the language model score simply indicates the probability of a given word, given a prior history of the word. The word match score simply scores the probability of a given word.

Neither of these scores has anything to do with a linguistic pattern, or a score for a linguistic pattern, much less a combination of a language model score, given a linguistic pattern and a translation model score given the linguistic pattern.

At no point does Berger et al. teach or suggest any incorporation of any input that is even similar to a linguistic pattern (much less a plurality of linguistic patterns), wherein the linguistic pattern represents a grouping of components relative to a phrase to be translated. The scores in Berger et al. simply have nothing to do with a linguistic pattern that represents groupings of components in the phrase to be translated.

Again, it is worth pointing out that the process described in the Berger et al. reference of generating a word match score and eventually a translation match score appears to be nothing more than use of a translation model to provide a probability of translation of a word in a first language into a word in a second language. It has nothing to do with the form of

a related linguistic pattern that represents objective groupings of components relative to the phrase to be translated. The cited reference simply neither teaches nor suggests such incorporation of linguistic pattern information into the determination of an overall score.

Therefore, Appellants submit that Berger et al. is completely missing the step of identifying a plurality of possible linguistic patterns and calculating a translation probability for the pattern based on a combination of a language model probability for the pattern and a translation model probability for the pattern. Appellants thus submit that claim 1 is not anticipated by Berger et al., and respectfully request the Board to reverse the Examiner's rejection of independent claim 1.

3. Claims 6, 7, 12, 14 and 15 are allowable.

Of this set of claims, claims 6 and 12 are independent claims and claims 7, 14 and 15 depend from the independent claims.

Both independent claims 6 and 12 include the step of "identifying a plurality of possible linguistic patterns..." discussed with respect to claim 1 above. However, both claims 6 and 12 also include calculating a translation probability for the phrase in the first language, into the second language, "in the linguistic pattern." In other words, these claims not only claim calculating a translation probability, but specifically claim that the translation probability is indicative of a probability of translating a phrase in the first language into a phrase in the second language that has the form of the linguistic pattern under analysis.

Therefore, not only is Berger et al. completely missing the step of identifying the possible linguistic patterns associated with a translation, as with respect to claim 1 above, but it is also completely missing the step of calculating a probability of translating the input phrase into a phrase in the second language which is in the form of the linguistic pattern.

Since Berger et al. is completely silent with respect to these elements of claims 6 and 12, Appellants submit that claims 6 and 12 are allowable over Berger et al. Therefore, Appellants respectfully request the Board to reverse the rejection of claims 6 and 12 as well as that of dependent claims 7, 14 and 15.

4. Claims 2, 3, 9, 10 and 16 are allowable.

Each of these claims include similar steps further defining their respective independent claims. By way of example, claim 2 includes the steps of "identifying a highest translation probability calculated; and identifying a linguistic pattern, for which the highest translation probability was calculated, as indicative of a likely phrase translation of the phrase in the first language."

Because Berger et al. is completely silent with respect to identifying linguistic patterns associated with a possible translation, Berger et al. cannot disclose any method of choosing a linguistic pattern as indicative of the proper translation of the input phrase. Therefore, Berger et al. does not teach or suggest identifying the linguistic pattern for which the highest translation probability was calculated as being indicative of the likely phrase translation of the input phrase. Thus, Berger et al. cannot teach or suggest the limitations set out in claims 2, 3, 9, 10 and 16. Because Berger et al. does not teach these limitations, Berger et al. does not anticipate the claims and therefore the claims are independently allowable over Berger et al. Appellants thus respectfully request that the Board reverse the rejection of claims 2, 3, 9, 10 and 16.

5. Claims 4, 11 and 13 are allowable.

Claims 4, 11 and 13 deal with the specific technique for incorporating linguistic patterns into a method or system for performing machine translation. Because Berger et al. completely fails to teach or suggest the notion of linguistic patterns incorporated into the translation model, it simply fails to teach

or suggest any specific way in which they could be incorporated.

By contrast, claim 4 specifically states that identifying a plurality of possible linguistic patterns includes "accessing a bilingual data store that includes linguistic patterns in the second language associated with the phrase in the first language." Similar limitations are found in claims 11 and 13.

Therefore, not only has Berger et al. completely failed to teach or suggest that linguistic patterns can be incorporated, there is no teaching, whatsoever, as to how those linguistic patterns could be incorporated into a translation model. These claims specifically state that they are incorporated by using a bilingual data store that has the linguistic patterns in the second language associated with the phrase in the first language. This is neither taught nor suggested by Berger et al. Therefore, Appellants respectfully request the Board to reverse the rejection of claims 4, 11, and 13.

6. Claims 5 and 8 are allowable over Berger et al.

Claims 5 and 8 specifically state that a pattern probability is calculated for the linguistic pattern itself. Not only does Berger et al. completely fail to teach or suggest identifying linguistic patterns, and using linguistic patterns in a translation calculation, or a translation model, but it specifically fails to teach or suggest how such linguistic patterns could be used. Therefore, there is no mention, whatsoever, of calculating a pattern probability corresponding to the linguistic patterns such that the overall translation probability can be calculated using the pattern probability. Because this is neither taught nor suggested by Berger et al., Berger et al. cannot anticipate claims 5 and 8 of the present application. Appellant thus respectfully requests that the Board reverse the rejection of claims 5 and 8.

7. Conclusion, claims 1-16 should be allowed.

In conclusion, Appellants respectfully submit that the present description clearly meets the written description requirement set out in 35 U.S.C. §112, first paragraph.

In addition, Berger et al. simply fails to teach or suggest all of the elements of the independent claims 1, 6, and 12. Additionally, a number of the dependent claims are independently allowable, and they are allowable by virtue of their dependency, either directly or indirectly, from allowable independent claims. Thus, Appellants respectfully request that the Board reverse the Examiner and find that claims 1-16 are in condition for allowance.

The Director is authorized to charge any fee deficiency required by this paper or credit any overpayment to Deposit Account No. 23-1123.

Respectfully submitted,

WESTMAN, CHAMPLIN & KELLY, P.A.

By: _____

Joseph R. Kelly, Reg. No. 34,847
Suite 1400 - International Centre
900 Second Avenue South
Minneapolis, Minnesota 55402-3319
Phone: (612) 334-3222 Fax: (612) 334-3312

JRK:slg

Appendix A

1.(Previously Amended) A computer-implemented method of processing a phrase in a first language for translation to a second language, comprising:

receiving the phrase in the first language;

identifying a plurality of possible linguistic patterns in the second language associated with the phrase in the first language, wherein each of the plurality of possible linguistic patterns represents a grouping of components relative to the phrase; and

for each pattern, calculating a translation probability for the pattern based on a combination of a language model probability for the pattern and a translation model probability for the pattern.

2. (Original) The method of claim 1 and further comprising:

identifying a highest translation probability calculated; and

identifying a linguistic pattern, for which the highest translation probability was calculated, as indicative of a likely phrase translation of the phrase in the first language.

3. (Original) The method of claim 2 and further comprising:

providing an output as a translation of the phrase in the first language to the second language based on the linguistic pattern identified.

4. (Original) The method of claim 1 wherein identifying a plurality of possible linguistic patterns, comprises:

accessing a bilingual data store that includes linguistic patterns in the second language associated with phrases in the first language.

5. (Original) The method of claim 1 wherein calculating a translation probability further comprises:

calculating a pattern probability for the pattern.

6. (Previously Amended) A computer-implemented method of processing a multi-word phrase in a first language for translation to a second language, comprising:

receiving the multi-word phrase in the first language;

identifying a plurality of possible linguistic patterns in the second language that correspond to the phrase in the first language, wherein each of the plurality of possible linguistic patterns represents a grouping of translation components relative to the phrase; and

calculating a translation probability for translation of the multi-word phrase in the first language to one of the plurality of linguistic patterns in the second language.

7. (Original) The method of claim 6 wherein calculating a translation probability comprises:

for each of the linguistic patterns identified, calculating the translation probability as a combination of a language model probability for the pattern in the second language and as a translation model probability for the phrase in the first language, given the linguistic pattern in the second language.

8. (Original) The method of claim 7 wherein calculating a translation probability further comprises:

calculating the translation probability based on a pattern probability for the linguistic pattern.

9. (Original) The method of claim 7 and further comprising:
identifying a highest translation probability calculated; and
identifying a linguistic pattern, for which the highest translation probability was calculated, as indicative of a likely phrase translation of the phrase in the first language.

10. (Original) The method of claim 9 and further comprising:
providing an output as a translation of the phrase in the first language to the second language based on the linguistic pattern identified.

11. (Original) The method of claim 7 wherein identifying a plurality of possible linguistic patterns, comprises:
accessing a bilingual data store that includes linguistic patterns in the second language associated with phrases in the first language.

12. (Previously Amended) A natural language processing system, comprising:
a pattern engine receiving a phrase in a first language and identifying a plurality of linguistic patterns in a second language, associated with the phrase in the first language, possibly corresponding to a translation of the phrase from the first language to the second language, wherein each of the plurality of linguistic patterns represents a grouping of components relative to the phrase; and
a probability generator configured to generate, for each linguistic pattern identified, a translation

probability for translating the phrase in the first language to the second language in the linguistic pattern.

13. (Original) The system of claim 12 wherein the pattern engine, comprises:

a bi-lingual data store storing phrases in the first language and corresponding linguistic patterns in the second language.

14. (Original) The system of claim 13 wherein the probability generator comprises:

a translation model, the probability generator being configured to generate the translation probability by accessing the translation model.

15.(Original) The system of claim 14 wherein the probability generator further comprises:

a language model in the second language, the probability generator being configured to generate the translation probability by accessing the language model.

16. (Original) The system of claim 15 wherein the probability generator is configured to:

identify a highest translation probability calculated; and

identify a linguistic pattern, for which the highest translation probability was calculated, as indicative of a likely phrase translation of the phrase in the first language.